A PALAEONTOLOGICAL DATA BASE AND ITS PROCESSING SYSTEM ON PERSONAL COMPUTER: A METHODOLOGICAL EXPERIMENT

by

E. MISZLIVECZ – G. TURCZI

(Department of Palaeontology, Eötvös University, H – 1083 Budapest, Kun Béla tér 2, Hungary

(Hungarian Geological Institute, H-1442 Budapest, P.O. Box 106, Hungary)

(Received: 7th May, 1986)

Abstract

The system we have developed may be used for storing and subsequently processing data relating to the taxa (specific marks, biometric values, references) offering the user an adaptable and constantly improvable method.

This system meets the following requirements:

- it enables the taxa's representing various taxonomic levels to be described;
- the references relating to the taxon may be stored;
- the taxa may be separated on the basis of the given taxonomic marks., with the references in the literature also accessible;
- it may be used together with cluster and with certain statistical programmes.

As a matter of course, the success of the work rests on the expertise of the palaeontologist, since the software may alter only the time factor and the number of the simultaneously accessible data.

An enormous amount of information is routinely collected by the palaeontologist on the investigated taxa. This data base is stored usually on different kind of cards, developed by the researcher himself. The determination of the taxa is based on this data base and on comparisons with published results. Studying the investigation process and the structure of information on taxa we have realized the necessity to develop a computerized data base operation system to help the palaeontologists in their work.

Our requirements relating this system are the following:

- it should make possible the discription of a given taxon on different – mostly generic and species – taxonomic levels;
- it should make available the references on publications necessary for the determination of the taxon;
- answers should be obtainable based on the systematic descriptions;
- statistical, cluster, etc. processing should be easily connected to the system.

The task is solved by two units, operating independently:

- a palaeontological system and
- a bibliographical system.

Input data of the palaeontological data base are the following:

- name of the taxon,

– synonym list of the taxon,
– morphological and biometrical features of the taxon, using a predetermined mask,
– description of the taxon in words in free format.

Concerning the above mentioned data types, the success of the data base depends mostly on the mask. Mask means a chart (data sheet, certificate), with references on the features of the taxon to be described. Special caution and care is needed to develop this description chart fulfilling all needs. Manipulation of a mask containing features of any taxon of a higher taxonomic category (class, family) is complicated and may put up debatable questions. According to our experiences the greatest problem of similar systems is the rigid chart (filling out mask) or the fixed code system. The taxon description contains subjective elements.

Keeping in mind these problems we decided to develop a versatile system, which can be further developed by the user, according to his needs and experiences. The user palaeontologist can make several mask types of generic, subgeneric or species group level, which can be expanded at any time.

Direct application of the data base is the multiple access to the data:
– direct access by taxon name;
– search based on given features by any types of mask (selection of all taxa sytisfying the given criteria);
– search based on key words and word groups in the description of the taxa;
– access to the references in the synonym list.

Input data of the bibliographical data base are the following:
– author;
– title;
– journal title or publisher of book
– key words (listing of all taxa discussed in the text; at palaeontological applications of the systems only)
– area for free use (for notes).

Access to and search for the stored information can be made according to the input types. The types (title, author, key words) are considered as free format texts, and can be searched for any word or group of words contained in them.

The most important character of the bibliographic system is the possibility of search according to key words, making available the literature references of a given topic (taxon). In the case of consequent, regular input, the programme provides standardized reference lists.

One of the most frequent tasks of a palaeontologist is to determine a taxon. In this case a search of the palaeontological system according to the mask types (based on observed characters of the fossil) provides the list of taxa fulfilling the desired criteria. Then a search of the bibliographical system according to key words provides the list of literature references of the taxon.

Taxonomic characters of a taxon INPUT

PALAEONTOLOGICAL SYSTEM

List of taxa OUTPUT
INPUT

BIBLIOGRAPHICAL SYSTEM

List of references OUTPUT

A data base filled with the necessary amount of data provides the possibilities for comparative taxonomic studies and/or statistical evaluations, which would last for an extremely long time by manual methods.

Naturally, the success depends on the expert palaeontologist, since the software can modify the time factor and the amount of operable data only.

Communication with the computer is as simple as possible and is highly flexible. The information can be given as free format text, and can be corrected or modified. The commands are given by pushbuttons.

Filling of the systems by data and testing is in progress. Our experiences with the programme will be published later.

Source programme: OS/L BASIC
Hardware: TPA/L–32
CM 5400 disk
printer.

REFERENCES

BRUNTON, C. H. C. (1979): The development of a computer based curatorial system for paleontology at the British Museum (Natural History). – Spec. Pap. Paleont., 22, 159–174, London

JONES, B. (1979): Data storage and retrieval for the paleontological collections, University of Alberta. – Spec. Pap. Paleont. 22, 175–188, London

SMITH, P. L. (1984): The implications of data base management systems to paleontology: a, discussion of Jurassic ammonoid data. (in press), Vancouver

WILEY, E. O. (1981): Phylogenetics: the theory and practice of phylogenetic studies. – Wiley Interscience Publications, New York.